PET Reconstruction With an Anatomical MRI Prior Using Parallel Level Sets

Matthias J. Ehrhardt*, Pawel Markiewicz, Maria Liljeroth, Anna Barnes, Ville Kolehmainen, John S. Duncan, Luis Pizarro, David Atkinson, Brian F. Hutton, *Senior Member, IEEE*, Sébastien Ourselin, Kris Thielemans, *Senior Member, IEEE*, and Simon R. Arridge

***Abstract*—The combination of positron emission tomography (PET) and magnetic resonance imaging (MRI) offers unique possibilities. In this paper we aim to exploit the high spatial resolution of MRI to enhance the reconstruction of simultaneously acquired PET data. We propose a new prior to incorporate structural side information into a maximum *a posteriori* reconstruction. The new prior combines the strengths of previously proposed priors for the same problem: it is very efficient in guiding the reconstruction at edges available from the side information and it reduces locally to edge-preserving total variation in the degenerate case when no structural information is available. In addition, this prior is segmentation-free, convex and no *a priori* assumptions are made on the correlation of edge directions of the PET and MRI images. We present results for a simulated brain phantom and for real data acquired by the Siemens Biograph mMR for a hardware phantom and a clinical scan. The results from simulations show that the new prior has a better trade-off between enhancing common anatomical boundaries and preserving unique features than several other priors. Moreover, it has a better mean absolute bias-to-mean standard deviation trade-off and yields reconstructions with superior relative ℓ^2-error and structural similarity index. These findings are underpinned by the real data results from a hardware phantom and a clinical patient confirming that the new prior is capable of promoting well-defined anatomical boundaries.**

***Index Terms*—Anatomical prior, magnetic resonance imaging, parallel level sets, positron emission tomography, total variation.**

Manuscript received December 17, 2015; revised March 11, 2016; accepted March 29, 2016. Date of current version August 30, 2016. This research was funded by the EPSRC (EP/K005278/1) and EP/H046410/1 and supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. M. J. Ehrhardt was supported by an IMPACT studentship funded jointly by Siemens and the UCL Faculty of Engineering Sciences. K. Thielemans and D. Atkinson are partially supported by the EPSRC grant EP/M022587/1. *Asterisk indicates corresponding author.*

*M. J. Ehrhardt was with the Centre for Medical Image Computing, WC1E 6BT London, U.K. He is now with the Department for Applied Mathematics and Theoretical Physics, University of Cambridge, CB3 0WA Cambridge, U.K. (e-mail: m.j.ehrhardt@damtp.cam.ac.uk).

P. Markiewicz, L. Pizarro, S. Ourselin, and S. R. Arridge are with the Centre for Medical Image Computing, WC1E 6BT London, U.K.

M. Liljeroth, A. Barnes, and K. Thielemans are with the Institute for Nuclear Medicine, University College London, NW1 2BU London, U.K.

V. Kolehmainen is with the Department of Applied Physics, University of Eastern Finland, 70211 Kuopio, Finland.

J. S. Duncan is with the Institute of Neurology, University College London, WC1N 3BG London, U.K.

D. Atkinson is with the Centre for Medical Imaging, University College London, NW1 2PG London, U.K.

B. F. Hutton is with the Institute for Nuclear Medicine, University College London, NW1 2BU London, U.K. and the Centre for Medical Radiation Physics, University of Wollongong, NSW, Australia.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2016.2549601

I. Introduction

POSITRON emission tomography (PET) allows monitoring with high sensitivity the distribution of a biologically important molecule and therefore to provide unique information for clinical applications; however, PET intrinsically suffers from low spatial resolution which, due to the partial volume effect, may prevent it from being quantitative [1]–[5]. High spatial resolution is one of the key strengths of magnetic resonance imaging (MRI) and is often available either from a separate scan with the help of registration [1]–[3], [6] or from a combined PET-MRI scanner that can simultaneously image function and structure [7]–[11].

The anatomical MRI information can be used to correct for the partial volume effect either post reconstruction [1]–[3], [5] or within the reconstruction [12]–[14]. Over the last two decades many priors have been proposed to utilize anatomical side information into the reconstruction of a low resolution modality [6], [15]–[30]. The proposed methods for this task often rely on a segmentation of the anatomical image [6], [15], are a heuristic modification of a minimization procedure [16], [17], [20] or minimize a non-convex functional [18], [21]–[24], [29], [30]. In all cases there is a compromise on stability, robustness and/or theoretical justification.

There have been priors proposed that do not rely on a segmentation and are convex [25], [26], [28], but these lack other desirable properties. It is important that a prior that incorporates anatomical information respects the information content in the functional image. As such, it is desirable that the prior reduces locally to an edge-preserved denoising scheme, such as total variation, if no *a priori* edge information is available; which is not the case for [26], [28]. Moreover, functional and anatomical images from PET and MRI might share many edges, but in general we cannot expect that the intensities change in the same way: at the edge of an anatomical region the MRI contrast might increase while the tracer uptake in PET might decrease or vice versa. This feature, although very important to combine images of arbitrary intensities, is not part of the model proposed in [25]. In this paper we combine the strengths of [25], [26] and propose a prior that does not rely on a segmentation, is convex, preserves the edges of unique features and does not rely on any assumptions on the intensities of the two images.

A. Contributions

The contributions in this paper are threefold. First, we propose a new prior to incorporate structural information into the reconstruction that has all the desired properties we outlined above and we prove its convexity. Second, we apply other priors

that have been used for other applications to the setting of PET-MRI. Finally, we compare five different priors in the setting of PET-MRI on synthetic phantom data, real phantom data and clinical patient data.

B. Set-Up

We consider PET data y as a random variable modelled as a Poisson process [31] with expectation

$$\mathbb{E}y = Au + r, \tag{1}$$

where $u\colon \Omega \to [0,\infty)$ denotes the PET image, A denotes the PET forward operator that includes scanner geometry, detector normalization and attenuation, and r denotes a background term needed to correct for scatter and randoms.

Based on this model, we perform image reconstruction via minimization of an objective function [32]

$$u^\sharp \in \operatorname*{argmin}_{u\geq 0} \Big\{ \mathcal{L}(Au + r, y) + \alpha \mathcal{R}(u) \Big\}, \tag{2}$$

where

$$\mathcal{L}(Au + r, y) := \sum_i (Au + r)_i - y_i \log[(Au + r)_i] \tag{3}$$

measures the distance of the estimated data $Au + r$ to the acquired data y. The data fit $\mathcal{L}$ is (up to an additive constant independent of u) the negative logarithm of the Poisson distribution which naturally calls for a non-negativity constraint on the image values $u(x)$ for $x \in \Omega$.

The prior $\mathcal{R}$ introduces *a priori* knowledge of the solution we seek. The regularization parameter α allows the balancing of information that comes from the data with our *a priori* belief about the solution $u^\sharp$. A popular prior is the (smooth) total variation [33]

$$\mathrm{TV}(u) := \int_\Omega \Big(\beta^2 + |\nabla u(x)|^2 \Big)^{1/2} \mathrm{d}x, \quad \beta > 0 \tag{4}$$

as it leads to edge-preserved denoising. Here β is introduced to render (4) differentiable; it is sometimes considered a scale parameter on the values of $|\nabla u(x)|$ below which edges are considered to be noise. We will return to a discussion of this parameter later.

In the context of PET-MRI, we have structural knowledge on the solution given by an anatomical MRI image $v\colon \Omega \to [0,\infty)$; we seek an extension of (4) which allows us to incorporate this *a priori* edge information. We will denote a regularization term that depends on an associate image by $\mathcal{R}(u|v)$. Moreover, we call the extra information about the structure *side information* which provides prior information about the PET image we seek aside the actual acquired data.

II. Methods

A. Asymmetric Parallel Level Sets

To simplify the notation, we introduce the spatially varying gradient field

$$\xi(x) := \frac{\nabla v(x)}{|\nabla v(x)|_\eta} \tag{5}$$

with the regularized norm $|x|_\eta := (|x|^2 + \eta^2)^{1/2}$, $\eta > 0$. The parameter η plays a similar role to β in (4) in that it scales down the influence of ∇v when edges merely represent noise. At any location x, the vector field $\xi(x)$ points in the direction of the gradient $\nabla v(x)$ but it is normalized such that

$$0 \leq |\xi(x)| \leq 1 \tag{6}$$

where the lower bound is obtained if $\nabla v = 0$ and the upper bound is obtained asymptotically as $|\nabla v| \to \infty$.

Motivated by the findings in [34]–[37] we can measure how structurally similar an image u is locally to another image v by comparing ∇u to the gradient field ξ by

$$0 \leq \Big(|\nabla u(x)|^2 - \langle \nabla u(x), \xi(x) \rangle^2 \Big)^{1/2} \leq |\nabla u(x)|, \tag{7}$$

where $\langle x, y \rangle := \sum_i x_i y_i$ denotes the Euclidean inner product. The upper bound is obtained when there is no structural side information, i.e., $\nabla v(x) = 0$, and the lower bound (asymptotically for $|\xi(x)| \to 1$) when $\nabla u(x)$ is aligned to $\xi(x)$ in the sense that there exists a $\lambda(x) \in \mathbb{R}$ such that $\nabla u(x) = \lambda(x)\xi(x)$. Note that for the case when $\nabla u = 0$, the gradient vector is aligned to any other vector ξ by allowing the parameter λ to be zero.

We derive a global prior by integrating this local measure of similarity over the entire domain

$$\mathcal{R}(u|v) = \int_\Omega \Big(|\nabla u(x)|^2 - \langle \nabla u(x), \xi(x) \rangle^2 \Big)^{1/2} \mathrm{d}x. \tag{8}$$

From the local properties it follows directly that this prior is non-negative and zero if and only if ∇u is aligned to ξ almost everywhere. As the gradient ∇u is perpendicular to the level sets of u, and ξ is perpendicular to the level sets of v, we refer to this measure of similarity of structures as the method of *parallel level sets*.

This prior has all the desired properties: it is convex in u, cf., [37] or proposition 1 in the appendix, it does not depend on a segmented MRI image and in the degenerate case when the MRI image is flat, i.e., $\nabla v(x) = 0$, it reduces to the total variation of the PET image. Moreover, it is independent of the sign and scale of ∇v and therefore can be applied to images of arbitrary intensities.

Similar to the case of total variation above, we introduce a smoothing parameter $\beta > 0$

$$\mathcal{P}(u|v) := \int_\Omega \Big(\beta^2 + |\nabla u(x)|^2 - \langle \nabla u(x), \xi(x) \rangle^2 \Big)^{1/2} \mathrm{d}x \tag{9}$$

that allows us to employ smooth minimization methods. The extension to the non-smooth case, i.e., $\beta = 0$, will be the subject of future work.

B. Other Methods to Incorporate Anatomical Information

We will benchmark our prior against previously proposed convex and segmentation-free priors [19], [25], [26], [28], [38], [39]. While some of these have been proposed for the very same application [19], [25], [28], others have been proposed for similar tasks in other applications such as geophysics [38] and colour imaging [39], other modalities like electrical impedance tomography (EIT) combined with computer assisted tomography [26] or joint PET-MRI reconstruction [35].

1) Kaipio et al.: Kaipio *et al.* proposed to incorporate *a priori* knowledge by the prior

$$\mathcal{K}(u|v) := \frac{1}{2}\int_\Omega |\nabla u(x)|^2 - \langle \nabla u(x), \xi(x)\rangle^2 \,\mathrm{d}x \qquad (10)$$

where ξ is defined as in (5) [26]. The original formulation is a little different but it is equivalent to (10) with a slightly different normalization, cf., [37] for details. This prior has most of the desired properties but, as can be readily seen, it reduces to a quadratic functional in the degenerate case $\xi \to 0$, rather than to total variation. This prior has been proposed originally in the EIT context and we apply it to the PET-MRI setting for the first time.

2) Kazantsev et al.: Motivated by the LOT model [40] and the Bregman distance for total variation [41], it has been proposed [25] to formulate the prior knowledge as

$$\mathcal{D}(u|v) := \int_\Omega \left(\beta^2 + |\nabla u(x)|^2\right)^{1/2} - \langle \nabla u(x), \xi(x)\rangle \,\mathrm{d}x, \quad (11)$$

where again a smoothing parameter is used to make the problem differentiable. This model fixes the problem of Kaipio *et al.* that features in u that are not present in v are penalized quadratically, and therefore allows edges in these areas. However, it penalizes the deviation of ∇u and ξ in a way so that vectors with opposite orientations are penalized even more than orthogonal vectors.

3) Bowsher's Prior: It has been proposed by Bowsher *et al.* to define a prior on neighbouring voxels by

$$\mathcal{B}(u|v) := \frac{1}{2}\sum_i \sum_{j\in N(i)} \omega_{i,j}(v)(u(i) - u(j))^2. \qquad (12)$$

The weights $\omega_{i,j}(v)$ are chosen such that the k most similar neighbours in the anatomical image have a positive weight depending on the spatial distance of voxel i to voxel j and zero otherwise. As it might happen that $u(i) - u(j)$ is weighted differently than $u(j) - u(i)$ we use a symmetrized version where the weights $\omega_{i,j}(v)$ and $\omega_{j,i}(v)$ are averaged.

4) Joint Total Variation: The last prior we benchmark against is joint total variation

$$\mathrm{TV}_\mathcal{J}(u|v) := \int_\Omega \left(\beta^2 + |\nabla u(x)|^2 + \gamma|\nabla v(x)|^2\right)^{1/2} \mathrm{d}x \quad (13)$$

where a parameter $\gamma > 0$ is used to adjust the scale of the side information. It has been first proposed as an extension of total variation to RGB colour imaging [39] and has subsequently been used for joint reconstruction in geophysics [38] and joint reconstruction of PET-MRI [35]. In contrast to $\mathcal{P}$, $\mathcal{K}$ and $\mathcal{D}$, joint total variation only makes use of the magnitude of the gradient of the side information thereby neglecting possibly valuable information. Recently a similar prior has been proposed to incorporate anatomical information into PET reconstruction [19].

An overview of the different methods with some key properties is given in Table I. Not all of the methods reduce to total variation in the degenerated case when no side information is available. While all of the methods depend on the location of the edges, only $\mathcal{D}$, $\mathcal{K}$ and $\mathcal{P}$ depend on the edge orientation. However, $\mathcal{D}$ does not allow edges to be negatively correlated. This means for example that if there is a "jump up" in the side information, then a "jump down" in the image to be reconstructed is strongly penalized. $\mathcal{P}$ is the only prior that fulfils all of the desired criteria.

TABLE I
SUMMARY OF ANATOMICAL PRIORS. THE LAST CATEGORY ONLY APPLIES FOR METHODS THAT ARE ORIENTATION DEPENDENT. $\mathcal{P}$ IS THE ONLY PRIOR THAT FULFILS ALL OF THE FOUR CRITERIA. † PROPOSED

	$\mathrm{TV}_\mathcal{J}$	$\mathcal{B}$	$\mathcal{D}$	$\mathcal{K}$	$\mathcal{P}^\dagger$
reduces to total variation	✓	✗	✓	✗	✓
edge location dependent	✓	✓	✓	✓	✓
edge orientation dependent	✗	✗	✓	✓	✓
allows negative edge correlation	-	-	✗	✓	✓

III. NUMERICAL SET-UP

Algorithm, Projections and Parameters

Algorithm: In order to fairly compare all the different priors, we use the same method to minimize (2) with the different choices of priors discussed in the last sections. To be more precise, we use L-BFGS-B [42], [43] where the non-negativity constraint is implemented by projecting the iterates onto the non-negative quadrant. L-BFGS-B is a Quasi-Newton method that approximates the inverse of the Hessian with first order information. In all cases, we run L-BFGS-B for 2000 iterations. Implementation is in MATLAB®.

As this paper focuses on priors rather than optimization algorithms we do not compare or investigate other algorithms. We plan to investigate optimization algorithms for this application more closely, particularly for the non-smooth case when $\beta = 0$.

1) Projections: All the data in this paper correspond to the geometry of one direct plane of the Siemens Biograph mMR®, cf., [44] for scanner specifications, fixed at a given axial position and formed by summing six or five (depending on the axial position) direct and cross sinograms in the scanner's native axial compression of span-11. The PET forward and adjoint operators for this geometry are taken from STIR (Software for Tomographic Image Reconstruction) [45] that has been interfaced to MATLAB®. In all cases we model the loss of resolution by a Gaussian blur of full width at half maximum (FWHM) of 4 mm $\times$ 4 mm in image space prior to projection and after backprojection.

2) Parameters: We tested several regularization parameters α for all methods and show a few of these results, cf., the model (2). The parameter η for $\mathcal{D}$, $\mathcal{K}$ and $\mathcal{P}$ has been chosen in $[10^{-3}, 10^{-2}]$ which is around 0.1%-1% of the maximal gradient magnitude of the side information. The similar parameter γ for $\mathrm{TV}_\mathcal{J}$ has been chosen in [1, 5] which leads to gradients of similar magnitude in both images. For the TV-like priors, we smoothed the norm by $\beta = 10^{-4}$ which is approximately 0.01% of the expected maximal gradient intensity of the PET image. For $\mathcal{B}$, 4 neighbours from a 3×3 neighbourhood were chosen. For comparison, we also ran maximum likelihood expectation maximization (MLEM) [46] for up to 500 iterations and smoothed the final iterate with a Gaussian filter with FWHM of 4 mm $\times$ 4 mm.

A. Phantoms

1) Software Phantom: The first test case is a software phantom, cf., Fig. 1(a), which is based on an MRI image

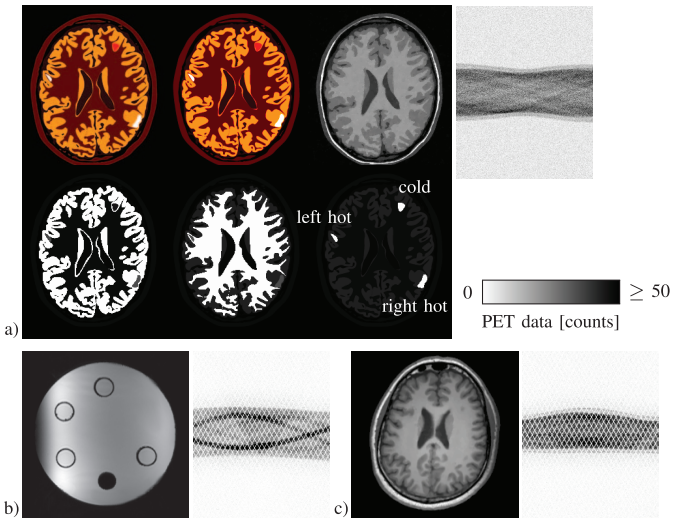

Fig. 1. MRI side information and PET data for (a) software phantom, (b) hardware phantom and (c) clinical patient data. For the software phantom also the PET ground truth (high resolution, low resolution) and regions of interest (grey matter, white matter, lesions) are shown.

obtained from BrainWeb [47] and converted into a continuous spline phantom [48]. Different regions in the brain such as grey matter, white matter, cerebrospinal fluid, cold lesion and hot lesions are then assigned a constant intensity reflecting an expected FDG uptake. The constant uptake in the regions has been modelled as 0.44, 0.11, 0.06, 0.28, 1, respectively. We sampled the continuous phantom on a resolution of 1140 × 1140 (0.25 mm × 0.25 mm). To simulate finite voxel sizes, the images are then averaged over 4 × 4 regions to get ground truth images of size 285 × 285 (1 mm × 1 mm). The noise level is set to 500 k counts with another 500 k counts contributing to the background. The randoms have been modelled spatially constant and the scatter smoothly varying, resembling the shape of the x-ray transform of the ground truth. In addition, known attenuation from a simulated CT is modelled as well. The regions of interests were set to be all pixels which contain at least 50% of a certain type (e.g., grey matter).

2) Hardware Phantom: For the second test case, shown in Fig. 1 b), a 6.4 litre cylindrical phantom was used for PET and MRI data acquisitions. The diameter and height of the phantom were 20.4 cm and 18.6 cm, respectively, with additional six inserts of the same size and diameter of 2.5 cm. One of the inserts was solid (made of Teflon). In order to obtain a good quality MRI signal, a solution of copper sulphate and sodium chloride was used in ratios of 4 g and 1 g, respectively, per one litre of water. This solution was further mixed with PET radiotracer, 18F-FDG, with varying radioactivity concentrations between the background and the inserts. The data was acquired on a Siemens Biograph mMR® hybrid PET-MRI scanner. For computational efficiency only the events detected in one direct compressed sinogram in span-11 were used. The sinogram corresponds to an axial position of 6.6 cm from the scanner's isocentre, directly covering the phantom's inserts. The sinogram plane was formed by summing six cross and uncompressed sinograms. The scatter and random events were estimated using the off-line version of the Siemens Healthcare reconstruction software.

3) Clinical Data: The clinical data are from a 34 year old, male epilepsy patient. The dataset is composed of a T_1-weighted MRI, a UTE-based μ-map and list mode FDG-PET data. The T_1-weighted MRI (3.0 T, TE: 2.63 ms, TR: 1700 ms, TI 900 ms, flip angle: 9°, voxel size: 0.53 × 0.53 × 1.1 mm), UTE-based μ-maps (voxel size: 1.56 × 1.56 × 1.56 mm) and PET list mode data (radiopharmaceutical: FDG) were acquired on a Siemens Biograph mMR® hybrid PET-MRI scanner; 250 MBq of FDG were administered half an hour before the 15 min PET acquisition. The MRI was co-registered to the PET image and then resampled using Vinci software [49] to account for motion between the MR and PET acquisitions. A slice of the MRI and the PET data acquired in one compressed sinogram in span-11 corresponding to one direct detection plane are shown in Fig. 1c).

IV. Results

A. Results for Software Phantom

1) Choice of Regularization Parameter: We will first investigate the choice of the regularization parameter α for the different methods. For MLEM the number of iterations can be seen as a regularization parameter. For noisy data the iterations exhibit a semi-convergence property such that we yield better results by early termination of the procedure [50]. To find a suitable choice we vary the regularization parameter and evaluate the result in terms of relative ℓ^2-error and structural similarity (SSIM) index [51], cf., Fig. 2. The optimal choice based on the relative ℓ^2-error over the whole phantom is marked in all four plots. We can see that for this choice of regularization parameter $\mathcal{K}$ and $\mathcal{P}$ perform best over the whole phantom for both quality measures (top row) and for grey matter (bottom left). However, this choice of regularization yields for $\mathcal{K}$ a suboptimal solution in the right hot lesion in terms of the relative ℓ^2-error, cf., right hand side of Fig. 2. This can also be seen in Fig. 3 and close-ups in Fig. 4 showing images for "optimal" regularization.

2) Perfect Versus Imperfect Side Information: Next, we compare reconstructions with perfect side information (namely the PET ground truth image) versus imperfect side information which is a reconstructed MRI image from noisy measurements, cf., Fig. 5. First of all, it should be noted that for perfect side information $\mathrm{TV}_{\mathcal{J}}$ and $\mathcal{B}$ result in larger errors compared to $\mathcal{K}$, $\mathcal{D}$ and $\mathcal{P}$. When the side information is changed to the more realistic MRI image, it can be seen that $\mathcal{D}$, in contrast to the other four methods, is not able to reconstruct the grey matter-to-white matter boundary as the side information is falsely informing that the activity should increase while in fact the activity decreases.

3) Bias-Versus-Standard Deviation Trade-Off: For $N = 50$ independent noise instances, we reconstruct PET images u_n, $n = 1, \dots, N$ and estimate with the mean $\mathbb{E} := (1/N) \sum_{n=1}^{N} u_n$, the *bias* and *standard deviation (SD)* as

$$\mathrm{bias} := \mathbb{E} - u^*, \quad \mathrm{SD} := \left(\frac{1}{N-1} \sum_{n=1}^{N} (\mathbb{E} - u_n)^2 \right)^{1/2}. \quad (14)$$

Here u^* denotes the PET ground truth that has been used for this simulation, cf., Fig. 1. The mean absolute value of these estimates over four regions of interest are shown in Fig. 6 as a

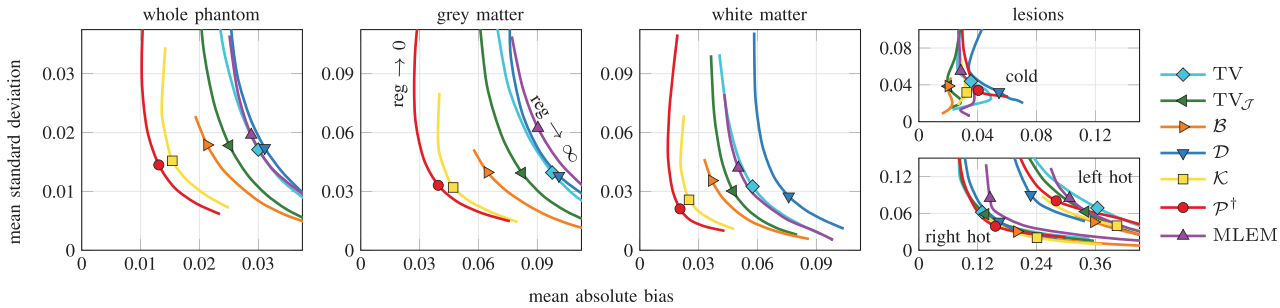

Fig. 2. Quantitative results for the software phantom with the amount of regularization (number of iterations for MLEM) on the horizontal axis. The optimal parameters are chosen based on the relative ℓ^2-error over the whole phantom (far left) and the results for this choice are marked solid in all four plots and shown in Fig. 3. Also shown (second left) the SSIM (structural similarity index) for the whole phantom. "Too low" and "too high" regularization, cf., Fig. 3, are marked with lighter shading. $\mathcal{K}$ and $\mathcal{P}$ perform best for the whole phantom and grey matter. Moreover, for this choice of regularization $\mathcal{B}$ and $\mathcal{K}$ perform worse for the right hot lesion than the other methods which all perform similarly. † proposed method.

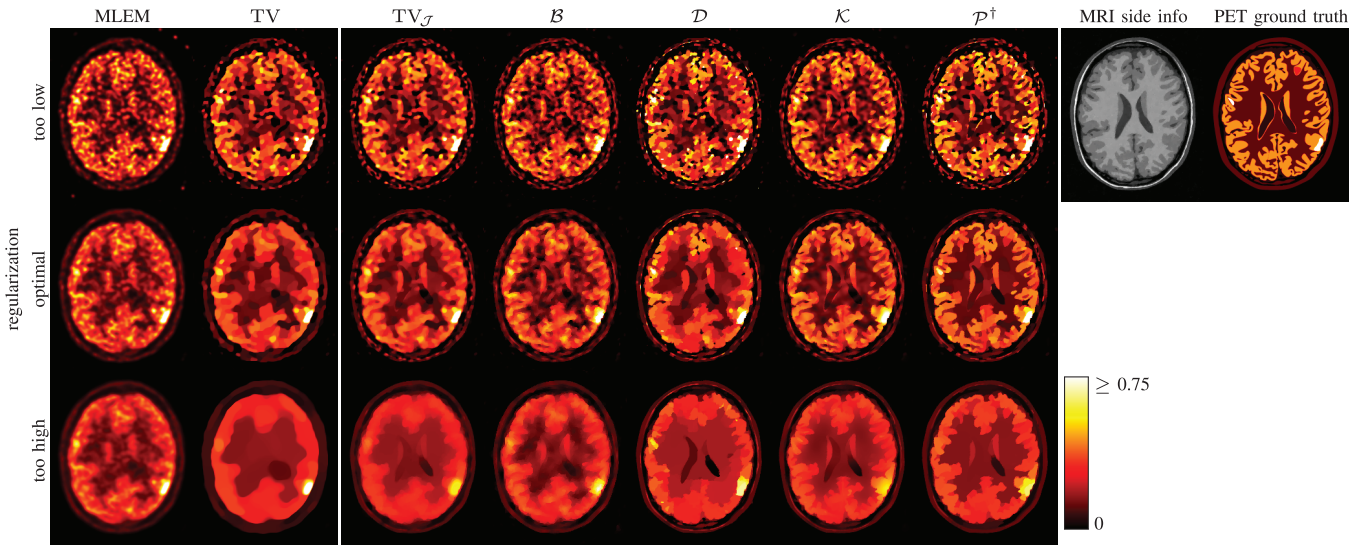

Fig. 3. Effect of regularization parameter (number of iterations for MLEM) on the results for the software phantom with MRI as side information. The choice is "optimal" with respect to the ℓ^2-norm between the reconstruction and the ground truth. The images correspond to the markers in Fig. 2. For the "optimal" choice of regularization, both $\mathcal{K}$ and $\mathcal{P}$ result in well-defined anatomical boundaries. † proposed method.

curve with respect to the regularization parameter/number of iterations. As it can be clearly seen, $\mathcal{P}$ closely followed by $\mathcal{K}$ has the best bias-versus-standard deviation trade-off for the whole phantom, grey matter and white matter. In the lesions, which are not present in MRI, all methods perform equally well. The "optimal" regularization parameter in terms of the expected mean squared error over the whole phantom is marked in all four plots. It can be seen that all methods have roughly the same standard deviation for the whole phantom, grey matter and white matter but $\mathcal{P}$ has always the smallest bias. In addition, while the methods that reduce to total variation (TV, TV$_\mathcal{J}$, $\mathcal{D}$, $\mathcal{P}$) have a similar bias and standard deviation in the right hot lesion, the two methods that reduce to a quadratic functional ($\mathcal{B}$, $\mathcal{K}$) have a slightly smaller standard deviation but larger bias in this region.

The bias and standard deviation for the regularization parameter with the smallest expected mean squared error are plotted as images in Fig. 7 with a line profile in Fig. 8. As it can be seen, $\mathcal{P}$ and $\mathcal{K}$ both have a smaller bias than the other five methods. In addition, we can clearly see again the bias of $\mathcal{D}$ at the grey matter-to-white matter interface. Moreover, the methods that reduce to total variation in absence of structural prior information TV, TV$_\mathcal{J}$, $\mathcal{D}$ and $\mathcal{P}$ have a more spatially localized standard deviation. In contrast, the standard deviation of $\mathcal{B}$ appears more spatially constant. Furthermore, we observe that the two small regions of higher activity at the right of the PET image are reconstructed with the least bias for $\mathcal{P}$ but have a higher standard deviation when compared for instance to TV. This effect is related to Bregman iterations that have been shown to *decrease the systematic bias* of total variation regularized reconstruction [41].

B. Results for Hardware Phantom

The results for the hardware phantom are shown in Fig. 9 with close-ups in Fig. 10 and line profiles in Fig. 11. Fig. 9 shows the results for all methods with a level of regularization chosen to balance data fitting accuracy and noise propagation. In order to minimize subjectivity, we also show images with lower and higher level of regularization. We would like to highlight three aspects that also correspond to the close-ups in Fig. 10. First of all, the hot insert that is not visible in the side information is, as expected, reconstructed well by the methods that reduce to total

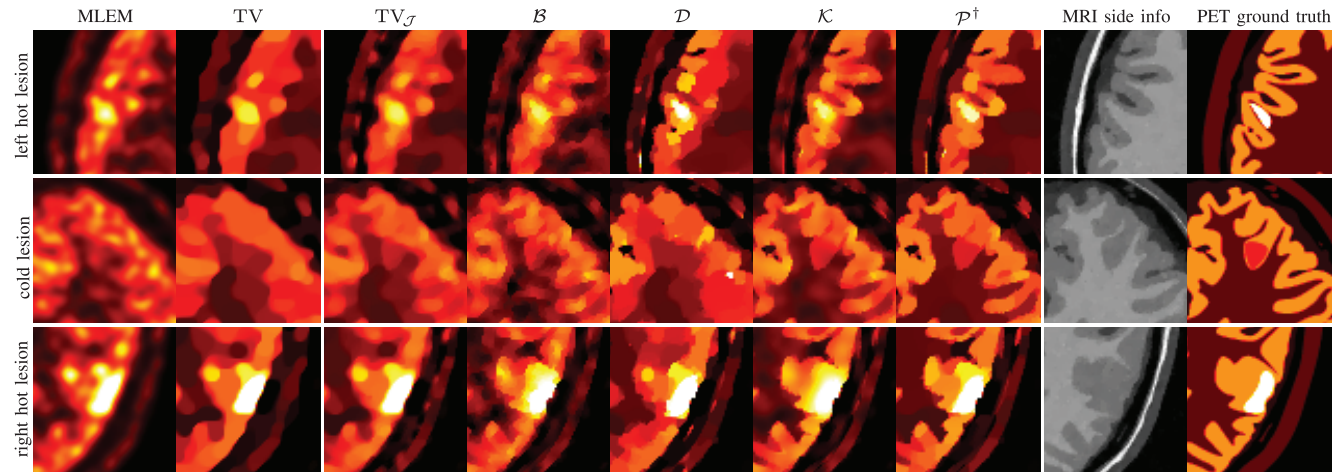

Fig. 4. Close-up of the lesions of the results with "optimal" chosen regularization parameter from Fig. 3. It can be seen that both $\mathcal{K}$ and $\mathcal{P}$ result in well-defined anatomical boundaries. Moreover, $\mathcal{P}$ shows clearly defined lesions.

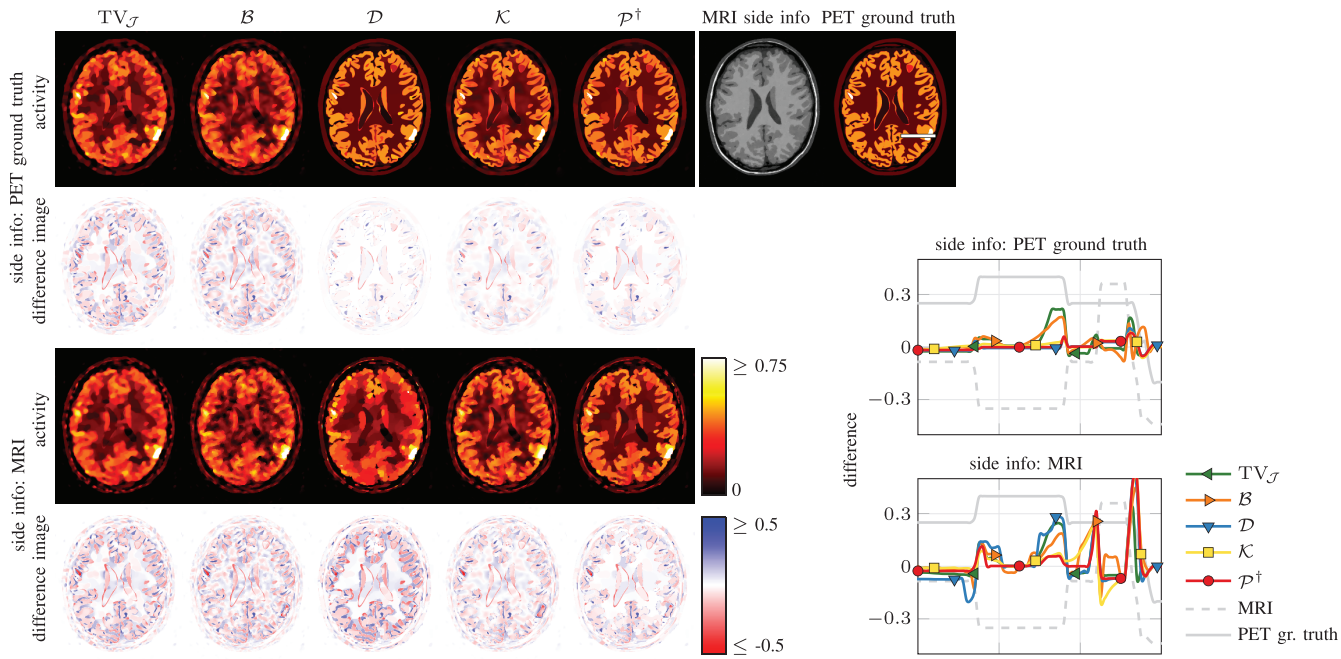

Fig. 5. Results for software phantom with ground truth of PET as side information at the top and reconstructed MRI at the bottom. Top: $\mathcal{D}$, $\mathcal{K}$ and $\mathcal{P}$ reconstruct the phantom almost perfectly with the PET ground truth given as side information. Bottom: $\mathcal{D}$ fails to reconstruct the grey matter-to-white matter interface as the gradients in PET and MRI are negatively correlated. The same observation can be made from the line profiles on the right where scaled and translated line profiles of the PET ground truth and the MRI serve as a reference. † proposed method.

variation while the other methods tend to over-smooth this feature. Second, at the hot insert at the bottom right we can see the same effect as at the grey matter-to-white matter interface of the software phantom: the prior $\mathcal{D}$ disfavours negatively correlated edges which results in a wide corona around the insert. Third, at the left edge of the phantom we can see that the intensity of the MRI phantom fades away. While it changes the smoothing behaviour of $TV_{\mathcal{J}}$, $\mathcal{B}$ and $\mathcal{K}$, it does not significantly affect $\mathcal{D}$ and $\mathcal{P}$. Finally, the line plots in Fig. 11 show that especially $\mathcal{P}$ results in clear, well-defined edges.

Note in Fig. 9, the sphere that appears in the PET reconstructions that are reconstructed with the MRI side information is not an artefact. It is clearly visible in the MLEM reconstruction with six times the number of counts. Thus, this sphere is not an artefact but shows that by using anatomical priors it is possible to detect an object with a very low contrast.

C. Results for Clinical Data

The results from the clinical patient data are shown in Fig. 12 with close-ups in Fig. 13. Although, we cannot say which result is the "best" we can make two observations. First, the resulting images for both $\mathcal{K}$ and $\mathcal{P}$ have well-defined anatomical boundaries with $\mathcal{K}$ superior in the level of detail. Second, as in the software phantom, $\mathcal{D}$ struggles to reconstruct the grey matter-to-white matter interface which appears very different compared to all other methods. The line profiles, shown in Fig. 14, confirm these observations. In addition, it

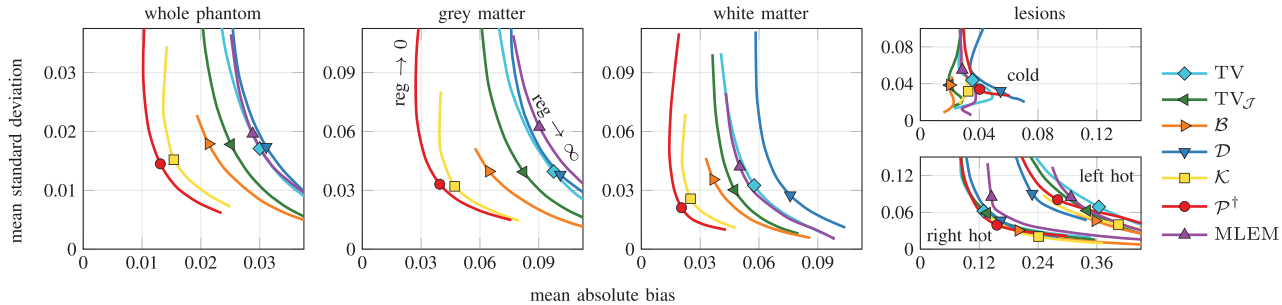

Fig. 6. Mean absolute bias-versus-mean standard deviation trade-off in different regions of interest. $\mathcal{K}$ and $\mathcal{P}$ have the best trade-off for the whole phantom, grey matter and white matter as their curves lie "underneath" the other curves but all methods have similar curves for the lesions (far right). Moreover, the solution that has the smallest expected mean squared error for the whole phantom (distance from the origin in the far left plot) is marked in all four graphs. It can be seen that these solutions have all roughly the same standard deviation for the whole phantom, grey matter and white matter but $\mathcal{P}$ has always the smallest bias. In addition, the "optimal" solution for $\mathcal{K}$ has a larger bias for the hot lesions.

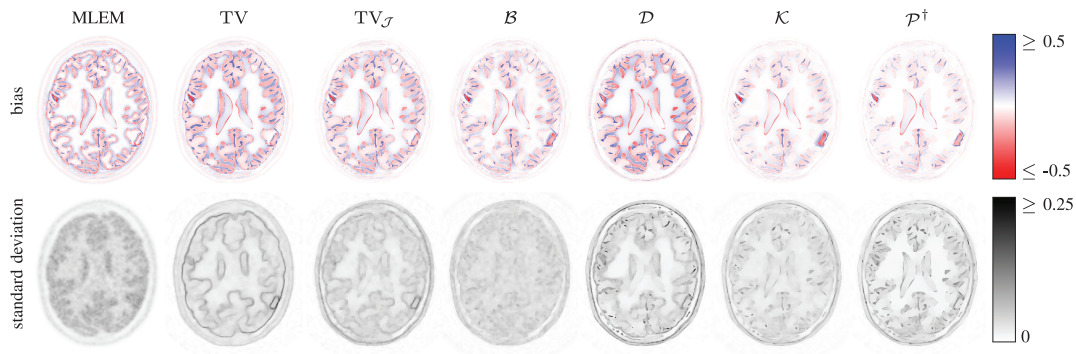

Fig. 7. Bias and standard deviation for software phantom for the regularization parameter that minimizes the expected mean squared error. $\mathcal{K}$ and $\mathcal{P}$ visually have the smallest bias with all methods appear to have a similar standard deviation. $\mathcal{D}$ shows a large bias at the grey matter-to-white matter interface.

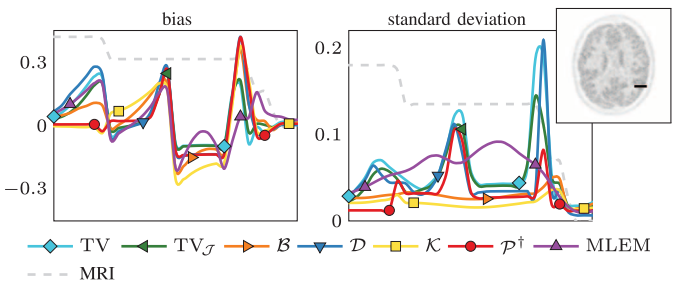

Fig. 8. Line profiles of bias and standard deviation for the software phantom, cf., Fig. 7. The line segment is marked in the image at the top right and a scaled version of the MRI serves as a reference about anatomical structure. All methods have spatially varying bias and standard deviation. While the methods that reduce to a quadratic prior, i.e., $\mathcal{B}$ and $\mathcal{K}$, have a relatively flat standard deviation, the standard deviation of the other three methods (which reduce to total variation in the absence of anatomical information) are relatively peaked.

can be seen that $\mathcal{P}$ results in a sharper hot spot than all other methods. However, no ground truth is available for this data.

V. Discussion

In this section we discuss prospects and limitations of anatomical priors.

A. Side Information

The results of this paper confirm that incorporating anatomical information can be very beneficial, for instance for recovering very low contrast features as in Fig. 9. However, to be useful, the reconstructed PET images have to be robust to errors in the anatomical information. Most priors (including the proposed $\mathcal{P}$) are insensitive to inhomogeneities in the MRI images, cf., Fig. 9.

The methodology relies on the registration of the two data sets which is intrinsically the case in the simultaneous PET-MRI set-up. However, even in this scenario we may encounter a slight misregistration—e.g., due to motion and distortions—which might introduce artefacts into the reconstruction. The sensitivity of anatomical priors in general, and the proposed method in particular, to such misregistration is out of the scope of this work but might be addressed in future research.

B. Parameters

This method has three important parameters that have to be chosen: The regularization parameters α, the edge parameter η and the smoothing parameter β. While we only show results for the selection of the regularization parameter α we briefly discuss our experience with the selection of the other two. First, the smoothing parameter β is related to the gradient magnitude that shall be preserved and can be chosen either from an unregularized MLEM reconstruction or based on previous reconstructions. We have found that the reconstructed images are not very sensitive to changes in this parameter up to at least a factor of ten. The edge parameter η should depend on the edge strength distribution of the anatomical image and defines which edges shall or shall not be encouraged in the PET image. The sensitivity to this parameter depends on the quality of the anatomical

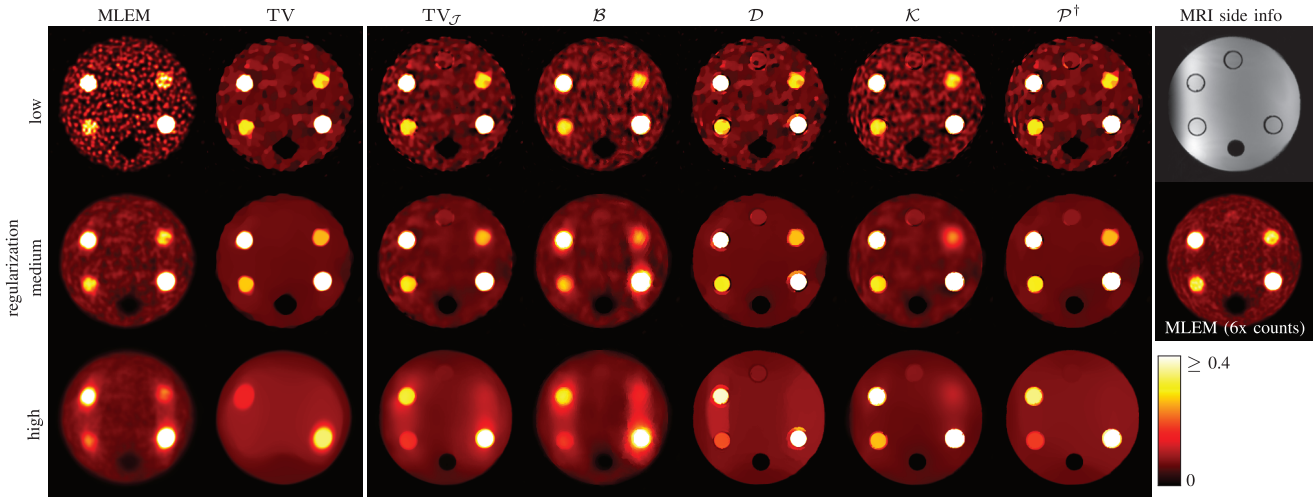

Fig. 9. Reconstructions of hardware phantom with MRI as side information for a varying amount of regularization. See caption of Fig. 10 for details.

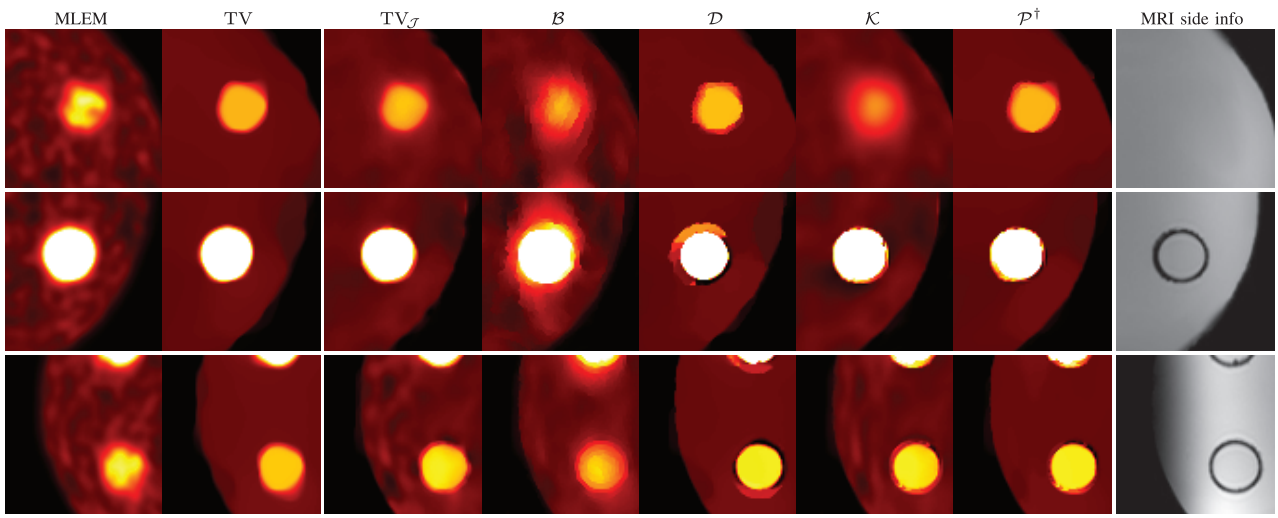

Fig. 10. Close-ups on inserts of the results for medium regularization (middle row) shown in Fig. 9. As it can be seen from the top row, $\mathrm{TV}_{\mathcal{J}}$, $\mathcal{D}$ and $\mathcal{P}$ do not smear out the hot insert that is not present in the side information. Method $\mathcal{D}$ does not allow negative gradient correlation and therefore introduces a corona around the inserts, cf., middle and bottom row. Both $\mathcal{D}$ and $\mathcal{P}$ reconstruct the left hand side of the phantom well despite the smooth variation in the side information (bottom row).

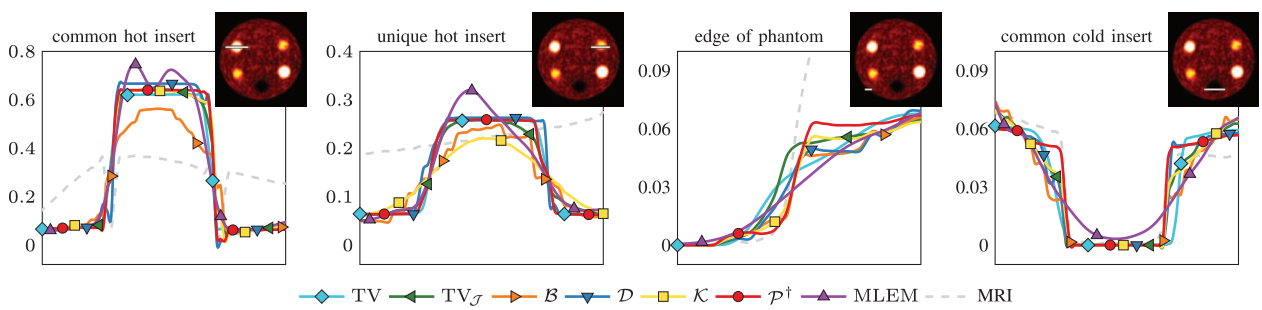

Fig. 11. Line segments of hardware phantom reconstructions for medium regularization shown in cf., Fig. 9 over several inserts and the edge of the phantom. The line segments are marked in the image at the right hand side. $\mathcal{D}$ and $\mathcal{P}$ yield the sharpest results without over-smoothing the insert not present in the side information (second left).

image as a small η can also encourage edges that are due to noise or other artefacts. Therefore, the scale of η is expected to be more important for images with less well-defined edges. As with all methods or algorithms we need to choose values of parameters that will influence the reconstructed image quality. In this paper we have chosen to optimize the parameters with respect to an objective and easily computable quality measure but it is important to note that "optimality" is very much application dependent and might need to involve humans that analyse the images. Therefore, all results should be interpreted with care

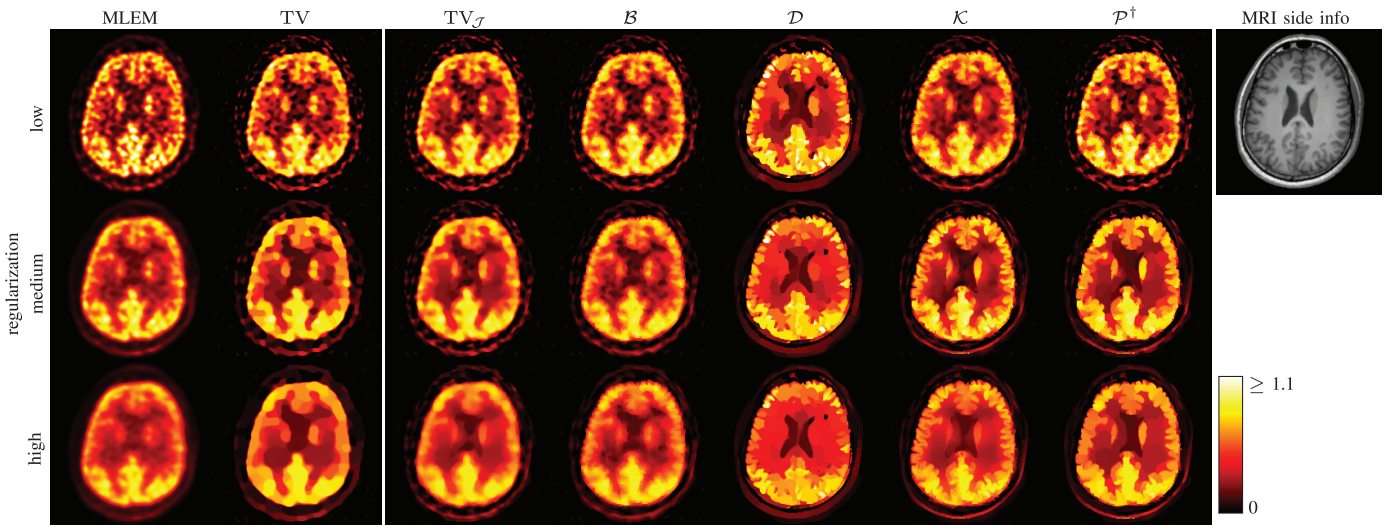

Fig. 12. Reconstructions of clinical patient data for a varying amount of regularization. It can be seen that both $\mathcal{K}$ and $\mathcal{P}$ result in image with well-defined anatomical boundaries. Moreover, $\mathcal{D}$ fails to reconstruct the grey matter-to-white matter boundary due to the negative correlation of the edge in PET and MRI. †proposed method.

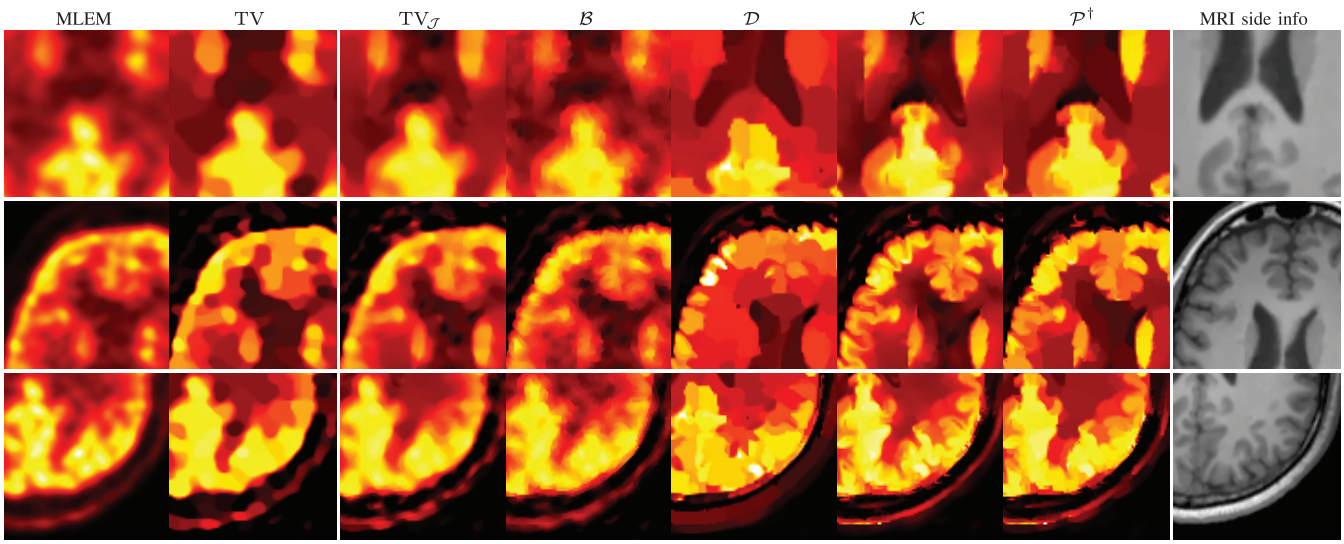

Fig. 13. Close-ups of results in Fig. 12 for medium regularization. Both $\mathcal{K}$ and $\mathcal{P}$ lead to well-defined structures. While $\mathcal{K}$ shows a higher level of detail in the grey matter, $\mathcal{P}$ shows a slightly sharper hot spot (see also Fig. 14 for the latter observation).

and a different parameter selection might be needed depending on the task.

C. Extension to 3D

All experiments in this paper have been carried out in two dimensions to save computation time. The extension to the three dimensional case might need more efficient algorithms that exploit all the structure of the problem. However, the mathematical basis is valid in arbitrary dimensions and there is no reason why this methodology should not translate to three dimensions. This will be the subject of future work.

VI. Conclusion

In this paper we proposed a new prior to incorporate structural side information into reconstruction and showed its application for the case of anatomical information from MRI incorporated into the reconstruction of PET. The proposed prior combines the strength of other previously published priors and has the advantage that it is convex, segmentation-free and edge-preserving in the degenerated case. The prior makes use of directional information from the anatomical side image and encourages images with aligned gradients or parallel level sets. Moreover, we introduced another prior that encourages parallel level sets—which reduces to a quadratic prior—to this particular application. Results from a simulated phantom, a hardware phantom and clinical data show that encouraging parallel level sets is very suitable for this application as it promotes well-defined edges and allows negative edge correlation. The proposed modification to combine the ideas of parallel level sets with total variation allows one to reconstruct distinct objects that are not present in the anatomical side information. The results for the software phantom show that the proposed prior is superior to the other tested priors in terms of several quality measures.

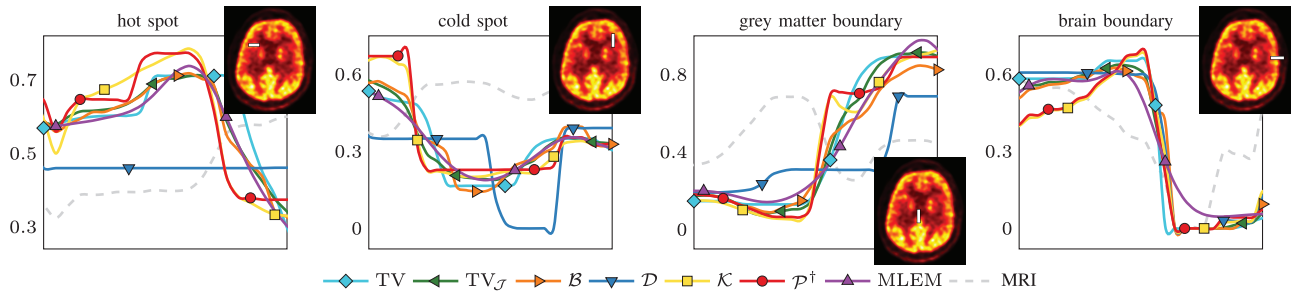

Fig. 14. Line segments of clinical data reconstructions for medium regularization, cf., Fig. 12. The line segments are marked in the image on the right. Both $\mathcal{K}$ and $\mathcal{P}$ yield similar results with sharp edges. In addition, $\mathcal{P}$ results in a well-defined hot spot (far left). The profiles of $\mathcal{D}$ do not match the profiles of the other methods apart from the edge of the brain (far right).

APPENDIX

The convexity of the proposed prior follows from arguments in [37]. We will state the proof for this special case here for completeness.

Proposition 1: The prior $\mathcal{P}$ defined in (9) *is convex.*

Proof: We will prove that at any location $x \in \Omega$ it holds

$$|\nabla u(x)|^2 - \langle \nabla u(x), \xi(x)\rangle^2 = |B_\xi(x)\nabla u(x)|^2 \qquad (15)$$

with $B_\xi(x)$ being a matrix which is independent of u. Therefore, the prior can be written as

$$\mathcal{P}(u|v) = \int_\Omega \left(\beta^2 + |B_\xi(x)\nabla u(x)|^2\right)^{1/2} \mathrm{d}x.$$

The convexity of $\mathcal{P}$ then follows from the convexity of $y \mapsto \sqrt{1+|By|^2}$ with a matrix B independent of y.

To prove (15), let $B_\xi := I - c\xi\xi^T$ with $c := (1+\sqrt{1-|\xi|^2})^{-1}$, where the spatial dependence on x has been omitted for readability. The latter is well-defined as $|\xi| \le 1$. We notice that c solves $2c - c^2|\xi|^2 = 1$ such that

$$\begin{aligned}|B_\xi \nabla u|^2 &= |\nabla u - c\langle \nabla u, \xi\rangle \xi|^2 \\ &= |\nabla u|^2 - 2\langle \nabla u, c\langle \nabla u, \xi\rangle\xi\rangle + |c\langle \nabla u, \xi\rangle \xi|^2 \\ &= |\nabla u|^2 - (2c - c^2|\xi|^2)\langle \nabla u, \xi\rangle^2 \\ &= |\nabla u|^2 - \langle \nabla u, \xi\rangle^2.\end{aligned}$$

■

REFERENCES

[1] H. W. Müller-Gärtner *et al.*, "Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects," *J. Cerbr. Blood Flow Metabolism*, vol. 12, no. 4, pp. 571–583, 1992.

[2] O. G. Rousset, Y. Ma, and A. C. Evans, "Correction for partial volume effect in PET: Principle and validation," *J. Nucl. Med.*, vol. 39, no. 5, 1998.

[3] C. C. Meltzer *et al.*, "Comparative evaluation of MR-based partial-volume correction schemes for PET," *J. Nucl. Med.*, vol. 40, no. 12, pp. 2053–2065, 1999.

[4] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-volume effect in PET tumor imaging," *J. Nucl. Med.*, vol. 48, no. 6, pp. 932–945, 2007.

[5] K. Erlandsson, I. Buvat, P. H. Pretorius, B. A. Thomas, and B. F. Hutton, "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology," *Phys. Med. Biol.*, vol. 57, no. 21, pp. R119–59, Nov. 2012.

[6] K. Baete, J. Nuyts, W. Van Paesschen, P. Suetens, and P. Dupont, "Anatomical-based FDG-PET reconstruction for the detection of hypometabolic regions in epilepsy," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 510–519, Apr. 2004.

[7] S. R. Cherry, "Multimodality in vivo imaging systems: Twice the power or double the trouble?," *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 35–62, 2006.

[8] D. W. Townsend, "Multimodality imaging of structure and function," *Phys. Med. Biol.*, vol. 53, no. 4, pp. R1–R39, 2008.

[9] M. S. Judenhofer *et al.*, "Simultaneous PET-MRI: A new approach for functional and morphological imaging," *Nat. Med.*, vol. 14, no. 4, pp. 459–65, Apr. 2008.

[10] B. J. Pichler, H. F. Wehrl, A. Kolb, and M. S. Judenhofer, "Positron emission tomography/magnetic resonance imaging: The next generation of multimodality imaging?," *Semin. Nucl. Med.*, vol. 38, no. 3, pp. 199–208, May 2008.

[11] C. Catana, A. R. Guimaraes, and B. R. Rosen, "PET and MR imaging: The odd couple or a match made in heaven?," *J. Nucl. Med.*, vol. 54, no. 5, pp. 815–24, May 2013.

[12] K. Vunckx *et al.*, "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 599–612, Mar. 2012.

[13] B. F. Hutton *et al.*, "What approach to brain partial volume correction is best for PET/MRI?," *Nucl. Instrum. Meth. A.*, vol. 702, pp. 29–33, Feb. 2013.

[14] B. Bai, Q. Li, and R. M. Leahy, "Magnetic resonance-guided positron emission tomography image reconstruction," *Semin. Nucl. Med.*, vol. 43, pp. 30–44, 2013.

[15] R. M. Leahy and X. Yan, "Incorporation of anatomical MR data for improved functional imaging with PET," in *Information Processing in Medical Imaging*. New York: Springer, 1991, LNCS, pp. 105–120.

[16] C. Chan, R. Fulton, D. D. Feng, and S. Meikle, "Regularized image reconstruction with an anatomically adaptive prior for positron emission tomography," *Phys. Med. Biol.*, vol. 54, no. 24, pp. 7379–400, 2009.

[17] D. Kazantsev *et al.*, "An anatomically driven anisotropic diffusion filtering method for 3D SPECT reconstruction," *Phys. Med. Biol.*, vol. 57, no. 12, pp. 3793–3810, 2012.

[18] S. Pedemonte, A. Bousse, B. F. Hutton, S. R. Arridge, and S. Ourselin, "Probabilistic graphical model of SPECT/MRI," in *Machine Learning in Medical Imaging*. New York: Springer, 2011, LNCS, pp. 167–174.

[19] L. Lu, J. Ma, Q. Feng, W. Chen, and A. Rahmim, "Anatomy-guided brain PET imaging incorporating a joint prior model," *Phys. Med. Biol.*, vol. 60, no. 6, pp. 2145–2166, 2015.

[20] K. Vunckx and J. Nuyts, "Heuristic modification of an anatomical Markov prior improves its performance," in *Proc. IEEE NSS-MIC*, 2010, pp. 3262–3266.

[21] J. Nuyts, "The use of mutual information and joint entropy for anatomical priors in emission tomography," in *Proc. IEEE NSS-MIC*, 2007, pp. 4149–4154.

[22] S. Somayajula, E. Asma, and R. M. Leahy, "PET image reconstruction using anatomical information through mutual information based priors," in *Proc. IEEE NSS-MIC*, 2005, pp. 2722–2726.

[23] S. Somayajula *et al.*, "Pet image reconstruction using information theoretic anatomical priors," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 537–549, Mar. 2011.

[24] J. Tang and A. Rahmim, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," *Phys. Med. Biol.*, vol. 54, no. 23, pp. 7063–75, 2009.

[25] D. Kazantsev, W. R. B. Lionheart, P. J. Withers, and P. D. Lee, "Multimodal image reconstruction using supplementary structural information in total variation regularization," *Sensing Imag.*, vol. 15, no. 1, p. 97, Jan. 2014.

[26] J. P. Kaipio, V. Kolehmainen, M. Vauhkonen, and E. Somersalo, "Inverse problems with structural prior information," *Inverse Probl.*, vol. 15, no. 3, pp. 713–729, 1999.
[27] D. Kazantsev, A. Bousse, S. Pedemonte, S. R. Arridge, and B. F. Hutton, "Edge preserving bowsher prior with nonlocal weighting for 3D SPECT reconstruction," in *Proc. IEEE NSS-MIC*, 2011, pp. 1158–1161.
[28] J. E. Bowsher *et al.*, "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," in *Proc. IEEE NSS-MIC*, 2004, pp. 2488–2492.
[29] J. Cheng-Liao and J. Qi, "PET image reconstruction with anatomical edge guided level set prior," *Phys. Med. Biol.*, vol. 56, pp. 6899–6918, 2011.
[30] J. Tang and A. Rahmim, "Anatomy assisted PET image reconstruction incorporating multi-resolution joint entropy," *Phys. Med. Biol.*, vol. 60, no. 1, pp. 31–48, 2015.
[31] J. M. Ollinger and J. A. Fessler, "Positron-emisson tomography," *IEEE Signal Process. Mag.*, vol. 14, no. 1, pp. 43–55, Jan. 1997.
[32] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, 2006.
[33] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1, pp. 259–268, 1992.
[34] M. J. Ehrhardt *et al.*, "Joint reconstruction of PET-MRI by parallel level sets," in *Proc. IEEE NSS-MIC*, 2014, pp. 1–6.
[35] M. J. Ehrhardt *et al.*, "Joint reconstruction of PET-MRI by exploiting structural similarity," *Inverse Probl.*, vol. 31, p. 015001, Jan. 2015.
[36] M. J. Ehrhardt and S. R. Arridge, "Vector-valued image processing by parallel level sets," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 9–18, Jan. 2014.
[37] M. J. Ehrhardt, "Joint reconstruction for multi-modality imaging with common structure," Ph.D. dissertation, Univ. College London, London, U.K., 2015.
[38] E. Haber and M. Holtzman-Gazit, "Model fusion and joint inversion," *Surv. Geophys.*, no. 34, pp. 675–695, Aug. 2013.
[39] G. Sapiro and D. L. Ringach, "Anisotropic diffusion of multivalued images with applications to color filtering," *IEEE Trans. Image Process.*, vol. 5, no. 11, pp. 1582–6, Nov. 1996.
[40] M. Lysaker, S. Osher, and X.-C. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1345–1357, Oct. 2004.
[41] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Sim.*, vol. 4, no. 2, pp. 460–489, 2005.
[42] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, no. 5, pp. 1190–1208, 1995.
[43] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York: Springer, 2000.
[44] G. Delso *et al.*, "Performance measurements of the siemens mMR integrated whole-body PET/MR scanner," *J. Nucl. Med.*, vol. 52, no. 12, pp. 1914–1922, 2011.
[45] K. Thielemans *et al.*, "STIR: Software for tomographic image reconstruction release 2," *Phys. Med. Biol.*, vol. 57, pp. 867–883, 2012.
[46] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. 1, no. 2, pp. 113–22, Oct. 1982.
[47] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, p. 425, 1997.
[48] M. Guerquin-Kern, L. Lejeune, K. P. Pruessmann, and M. Unser, "Realistic analytical phantoms for parallel magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 626–636, Mar. 2012.
[49] S. Vollmar *et al.*, "Vinci: Volume imaging in neurological research, co-registration and ROIs included," *Res. Sci. Comput.*, pp. 115–131, 2003.
[50] D. L. Bailey, D. W. Townsend, P. E. Valk, and M. N. Maisey, *Positron Emission Tomography – Basic Sciences*. New York: Springer, 2005.
[51] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–12, Apr. 2004.